EDGEWOOD CHEMICAL BIOLOGICAL CENTER
U.S. ARMY RESEARCH, DEVELOPMENT AND ENGINEERING COMMAND
Aberdeen Proving Ground, MD 21010-5424

ECBC-TN-068

# MODIFIED MAXIMUM LIKELIHOOD ESTIMATION METHOD FOR COMPLETELY SEPARATED AND QUASI-COMPLETELY SEPARATED DATA FOR A DOSE-RESPONSE MODEL

**Kyong H. Park**
**Steven J. Lagan**

**RESEARCH AND TECHNOLOGY DIRECTORATE**

**August 2015**

Disclaimer

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorizing documents.

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



| 1. REPORT DATE *(DD-MM-YYYY)* | 2. REPORT TYPE | 3. DATES COVERED *(From - To)* |
|---|---|---|
| XX-08-2015 | Final | Jun 2014 – Aug 2014 |

**4. TITLE AND SUBTITLE**
Modified Maximum Likelihood Estimation Method for Completely Separated and Quasi-Completely Separated Data for a Dose-Response Model

**5a. CONTRACT NUMBER**

**5b. GRANT NUMBER**

**5c. PROGRAM ELEMENT NUMBER**

**6. AUTHOR(S)**
Park, Kyong H.; and Lagan, Steven J.

**5d. PROJECT NUMBER**
CB10128

**5e. TASK NUMBER**

**5f. WORK UNIT NUMBER**

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Director, ECBC, ATTN: RDCB-DRI-M, APG, MD 21010-5424

**8. PERFORMING ORGANIZATION REPORT NUMBER**
ECBC-TN-068

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
Defense Threat Reduction Agency, 8725 John J. Kingman Road, MSC 6201, Fort Belvoir, VA 22060-6201

**10. SPONSOR/MONITOR'S ACRONYM(S)**
DTRA

**11. SPONSOR/MONITOR'S REPORT NUMBER(S)**

**12. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for public release; distribution is unlimited.

**13. SUPPLEMENTARY NOTES**

**14. ABSTRACT:**

When data are completely or quasi-completely separated, the traditional maximum likelihood estimation (MLE) method generates infinite estimates. The bias-reduction (BR) method, which is a variant of the bias-correction method, removes the first-order bias term by applying a modified score function, and it always produces finite estimates. By comparison, the traditional MLE method is unreliable because it produces infinite estimates. Unlike the Bayesian method, which may be difficult to apply in some situations, the BR method does not require prior information. The purpose of this paper is to provide all of the necessary equations and procedures required to carry out the BR method, thereby enabling use of the method on common computing platforms such as Basic in Microsoft Excel and Minitab. The U.S. Army Edgewood Chemical Biological Center members have been using related probit slopes for separated data, but such slopes are not always known. The BR method is a useful alternative for estimating parameters of separated data sets and small-sample sizes.

**15. SUBJECT TERMS**

| | | |
|---|---|---|
| Modified score function | Logit | Probit |
| Bias-reduction (BR) method | Bias-correction (BC) method | Bayesian method |
| Traditional maximum likelihood estimation (MLE) method | | R software |

| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT | 18. NUMBER OF PAGES | 19a. NAME OF RESPONSIBLE PERSON |
|---|---|---|---|---|---|
| a. REPORT | b. ABSTRACT | c. THIS PAGE | | | Renu B. Rastogi |
| U | U | U | UU | 30 | 19b. TELEPHONE NUMBER *(include area code)* (410) 436-7545 |

Blank

## EXECUTIVE SUMMARY

Applying a generalized linear model (GLM) with a logit or probit link is a routine procedure for estimating the effective or lethal dose for a dose-response model. The traditional maximum likelihood estimation (MLE) method used in GLMs generates infinite estimates whether data are completely separated (CS) or quasi-completely separated ([quasi-CS] i.e., when the range of doses for one of the responses [0 or 1] does not overlap the range of doses for the other response, or when it overlaps at only a single-dose level).

The bias-reduction (BR) method, described in this paper, removes the first-order bias term by applying a modified score function. This method always generates finite estimates, which embody all of the properties inherent in traditional MLEs. Unlike the Bayesian method, which may not be practical in many different situations, the BR method does not need prior information.

The purpose of this paper is to provide all the necessary formulas and procedures to carry out the BR method to generate finite estimates. This method could be used in common computing platforms (e.g., Microsoft Excel and Minitab) that are not pre-programmed to execute the BR method. U.S. Army Edgewood Chemical Biological Center members have been using related probit slopes when dealing with separated data, but the BR method would provide an additional option in cases where the probit slopes are not known. Additionally, the BR method can be used in small samples to reduce first-order asymptotic bias of parameter estimates. The BR method was written in R software (listed in Appendix B), and parameter estimates were compared between the BR and traditional MLE methods for three different data sets (i.e., CS, quasi-CS, and overlapped). The BR method produced finite confidence intervals (CIs) for all three data sets, whereas the traditional MLE method generated infinite CIs for the two separated data sets.

Blank

## PREFACE

The work described in this report was authorized under project number CB10128. The work was started in June 2014 and completed in August 2014.

The use of either trade or manufacturers' names in this report does not constitute an official endorsement of any commercial products. This report may not be cited for purposes of advertisement.


## Acknowledgments

The authors thank Mike Kierzewski for supporting this program.

Blank

## CONTENTS

## FIGURES



## TABLES

# MODIFIED MAXIMUM LIKELIHOOD ESTIMATION METHOD FOR COMPLETELY SEPARATED AND QUASI-COMPLETELY SEPARATED DATA FOR A DOSE-RESPONSE MODEL

## 1. INTRODUCTION

Maximum likelihood estimation (MLE) for a generalized linear model (GLM) is the most widely used method for estimating parameters for dose-response (DR) experiments with binary random 0 and 1 response variables. For DR models, a log-likelihood binomial distribution, with a logit or probit link, is a default method in statistical software for estimating the 50$^{th}$ percentile of the lethal dose (LD50) or the 50$^{th}$ percentile of the effective dose (ED50) for risk estimates.

The MLE method does not apply when data are completely separated (CS) or quasi-completely separated (quasi-CS) between binary responses (i.e., when the range of doses for the responses [0 or 1] does not overlap or overlaps at only a single-dose level). The conditions of existence for MLE estimates are discussed in several papers,[1,2] and estimates from the MLE method go to infinity for the CS or quasi-CS binary data. Infinite estimates can be considered as inaccurate with infinite confidence intervals (CIs). Unique and finite estimates exist when data are overlapped in explanatory variables for binary responses. Graphs with the three different data distributions are shown for CS doses (Figure 1), quasi-CS doses (Figure 2), and overlapped doses (Figure 3).

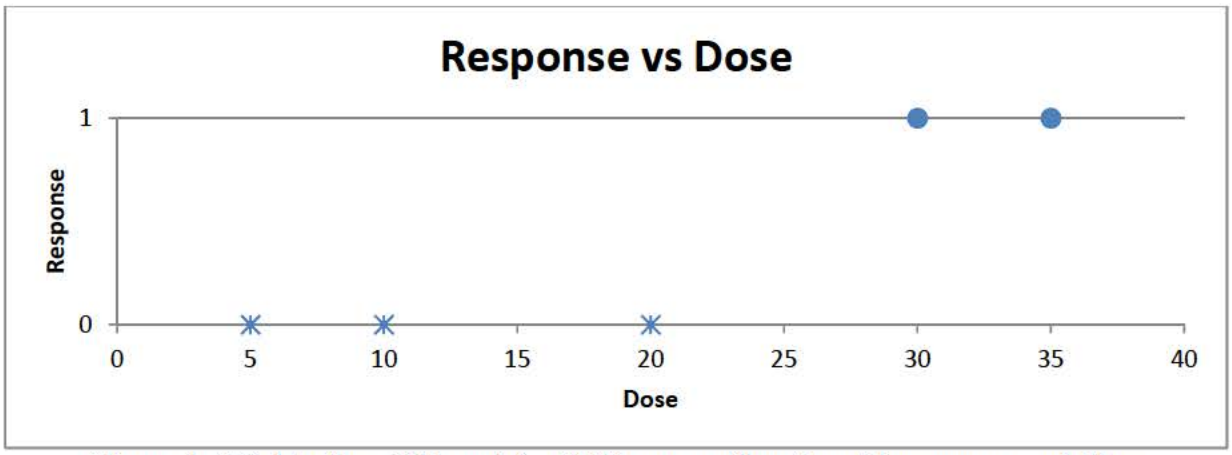


Figure 1. CS data for a DR model with binary random 0 and 1 response variables.

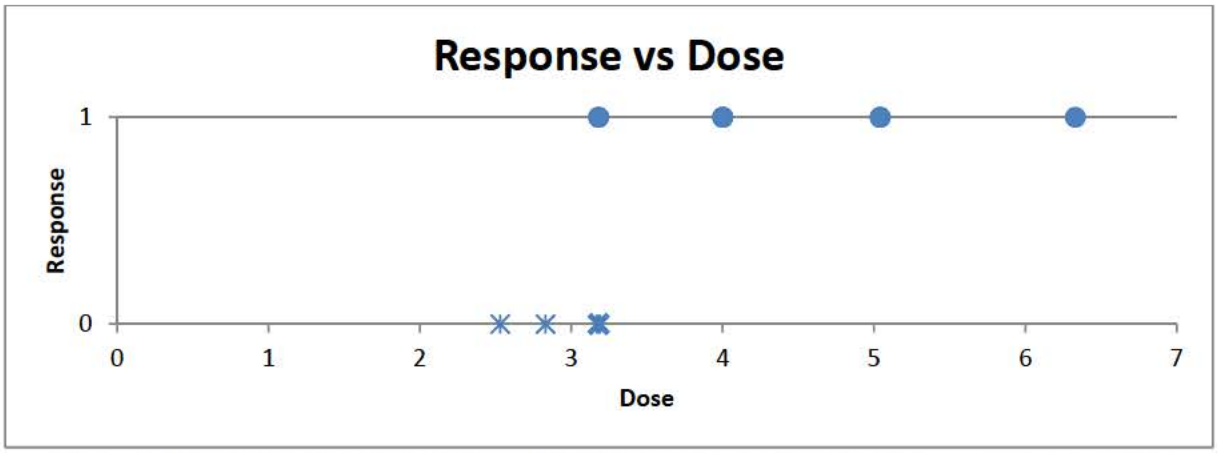


Figure 2. Quasi-CS data for a DR model with binary random 0 and 1 response variables (two responses overlap at dose level 3.18).

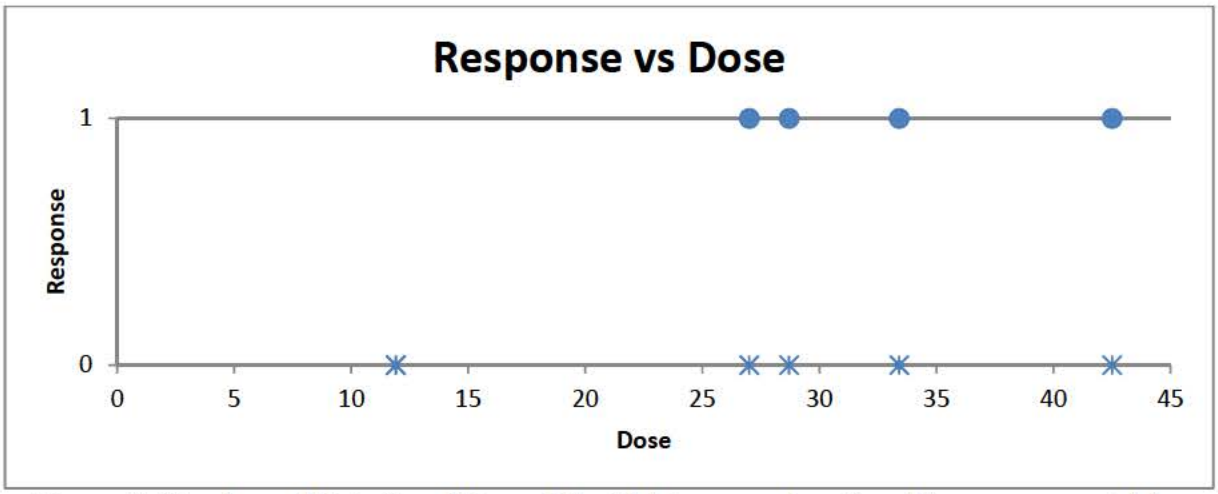


Figure 3. Overlapped data for a DR model with binary random 0 and 1 response variables (two responses overlap at several dose levels).

The MLE method is used to estimate the parameters of a nonlinear regression from score equations $\frac{\partial logL}{\partial \beta_i} \equiv U(\beta_i) = 0$ (Section 2) where log$L$ is a log-likelihood function. Firth recommends a first-order bias-reduction (BR) method for small-sample sizes based on a modified score equation.[3] Firth's modified MLE (MMLE) function is related to the penalized log-likelihood (PL) function, and its formulation was suggested as a solution for the CS and quasi-CS data. This equation will be discussed in detail in Section 2.

Other methods for overcoming the problems posed by the CS and quasi-CS data prevail, such as incorporating prior information to augment or better estimate the current data.[4] However, providing previously collected data is not always practical or possible.

When calculating LD50 or ED50 values for chemical warfare and toxic industrial agent human risk estimates at U.S. Army Edgewood Chemical Biological Center (ECBC), probit links (i.e., statistical relationships) have traditionally been used to perform analyses. Probit

slopes and LD50 or ED50 values are necessary parameters in calculating casualty estimates for downwind hazardous prediction models such as VLSTrack[a] and HPAC.[b] Whenever possible, ECBC members have been using previously known, related probit slopes for the CS or quasi-CS data sets.

Most of the published papers that were researched for this report discussed the MMLE method for the exponential family with a canonical link (EFCL) such as logistic binomial distribution. Because not much information is available on the necessary formulas for estimating CS and quasi-CS data-probit-link parameters, the purpose of this paper is to provide all the necessary formulas and procedures needed for estimation of finite parameters. These formulas may widen the application of the method to the other platforms that do not have built-in modules to carry out MMLEs.

Section 2 discusses MLE properties and bias-correction (BC) derivations of first-order asymptotic MLE bias. The BR method used for applying modified score functions and their relationships to PL functions for the logistic binomial function are also discussed. The differences between the BC and BR methods are also briefly discussed, and general equations for the BC and BR methods are provided. For the BR method, two different formulas are provided for canonical and non-canonical links. Section 3 provides the estimation process for applying an iterative reweighted procedure and obtaining initial estimates to begin the procedure. Section 4 summarizes the methods and procedures for the BR method. The R codes ([R Programming Language and Software] Ross Ihaka and Robert Gentleman, University of Auckland, New Zealand), which are used to estimate an intercept and a slope using the BR method, are listed in Appendix B.

## 2. BACKGROUND FOR DERIVATION OF MODIFIED SCORE FUNCTION

### 2.1 Properties of MLE Values in GLM

Generally, GLM is composed of three components: (1) The random component of response values, *y*; (2) the systematic component producing a linear predictor, $\theta = \sum_1^p x_i \beta_i$; and (3) a link function between the random and systematic components, $\theta = g(\mu) = X\beta$, where $g(\mu)$ is a link function. The random component, *y*, with the exponential family (such as binomial distribution) has the following distributional form:[6]

$$f = \exp\left\{\frac{(y\theta - b(\theta))}{a(\emptyset)} + c(y, \emptyset)\right\} \quad (1)$$

with distribution-specific functions a(.), b(.), and c(.). For the binomial distribution, the dispersion parameter is $a(\varphi) = 1$, and it can be expressed in the log form of the exponential formula:

[a] VLSTrack is Vapor, Liquid, and Solid Tracking.
[b] HPAC is Hazard Prediction and Assessment Capability.

$$\log(f) = y\log(\frac{\pi}{(1-\pi)}) - n\log(1-\pi)^{-1} + \log(\binom{n}{y}) \quad (2)$$

The canonical link derived for logit from eq 2 is expressed as

$$\theta = \log(\frac{\pi}{(1-\pi)}) \quad (3)$$

The logit model fits to the Θ for a linear model and transforms back to π. For probit, the link is

$$\theta = F^{-1}(\pi) \quad (4)$$

where $F(x_i^{'}\beta)$ is a standard cumulative normal distribution expressed as

$$F(x_i'\beta) = \int_{-\infty}^{x_j'\beta} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy \quad (5)$$

The MLE method is used to estimate regression parameters $\beta_j$ ($j$ = 1,…,$p$) by maximizing the likelihood function used to apply first-order derivatives with regard to parameter $\beta_i$. The log-likelihood function for the probit model linked with the binomial distribution is

$$\ln L = \sum_{i=1}^{n} \left\{ y_i \ln \frac{F(x_j'\beta)}{1 - F(x_j'\beta)} + \ln(1 - F(x_j'\beta)) \right\} \quad (6)$$

where $y_i$ is response variables 0 and 1. The first-order derivatives of the above log-likelihood function are

$$\frac{\partial \ln(L(\beta))}{\partial \beta_j} = \sum_{i=1}^{n} \frac{f(x_j'\beta)}{F(x_j'\beta)\left(1 - F(x_j'\beta)\right)} (y_i - F(x_j'\beta)) x_i \quad (7)$$

where $j$ is the $j^{\text{th}}$ explanatory variable, and $f(x)$ denotes a standard normal-density function

$$f(x) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\left(\frac{x^2}{2}\right)} \quad (8)$$

The estimates are calculated leading to the above score equations as

$$\frac{\partial \ln L(\beta)}{\partial \beta_j} = U(\beta_j) = 0 \quad (9)$$

The MLE method has many desirable properties for large samples. For example, MLE $\tilde{\beta}$ is a consistent estimator of true parameter $\beta_0$ with the probability approaching 1 as $n \to \infty$, that is, $\tilde{\beta}_{\text{mle}} \xrightarrow{P} \beta_0$. The consistent solution is asymptotically normal,

$\frac{\sqrt{n}}{\sigma}\left(\tilde{\beta}-\beta_0\right) \xrightarrow{L} Z$ (i.e., $\beta_{\text{mle}}$ converges to distribution $Z$, where $Z$ is a standard normal distribution, and asymptotic variances of unbiased estimators follow $I(\beta_0)^{-1}$, where $I(\beta_0)$ is the Fisher information matrix). The information matrix for the probit model is expressed as

$$I(\beta) = E(\frac{\sigma x_j'\beta}{\sigma\beta})^2 \;=\; -E\left(\frac{\sigma^2 \log\left(x_j'\beta\right)}{\sigma\beta^2}\right) = \int_{i=1}^{n} \frac{f^2(x_j'\beta)}{F\left(x_j'\beta\right)(1-F\left(x_j'\beta\right))} X'X \tag{10}$$

where $X'X$ is an explanatory matrix including a vector of 1 for an intercept. The MLE estimates are asymptotically unbiased and reach the Cramer-Rao bound,[7] which may be expressed as

$$\text{var}\left(\tilde{\beta}\right) \geq \frac{1}{I(\beta_0)} \tag{11}$$

## 2.2 BC Method Derivation

For CS and quasi-CS data, the estimates from the above score equations go to infinity with large variances exceeding thousands, which makes the estimates inaccurate with CIs reaching infinity. To remedy the infinite estimates, the score function is modified when derived from the BC method.

The score function computed from the profile likelihood is biased[3] because of the combination of the curvature and unbiased nature of the score function (i.e., 0 expectation of the score function expressed as $E\{U(\beta_0)\} = 0$ at the true value of $\beta$).

The expected estimators of MLE vector $\beta$s may be expressed as $E\left(\tilde{\beta}\right) = \beta_0 + \frac{b_1(\beta_0)}{n} + \frac{b_2(\beta_0)}{n^2} + \frac{b_3(\beta_0)}{n^3} +$ ..., where $\beta_0$ is the true unknown parameter; $\beta_i$ is the bounded function of $\beta_0$; and $n$ is usually the number of observations. BC is generally accomplished in two steps:
(1) maximum likelihood estimates are calculated and (2) the estimates are corrected up to the first-order bias term by subtracting it from the asymptotic bias of the maximum likelihood estimates of $\tilde{\beta}$ (i.e., $\tilde{\beta}_{\text{BC}} = \tilde{\beta} - \frac{b_1(\beta_0)}{n}$).

Many authors discussed the BC technique subtraction of the first-order bias term, especially for distributions in the EFCL, such as the logistic binomial function.[8,9] Nelder provided a formula for the first bias term, $\frac{b_1(\beta_0)}{n}$, for the EFCL, which was expressed as

$$\frac{b_1(\beta_0)}{n} = (X^T W X)^{-1} X^T W \epsilon \tag{12}$$

where

$$\epsilon_i = -\frac{1}{2} Q_{ii} \frac{k_{3i}}{k_{2i}} \tag{13}$$

and $Q_{ii}$ and $W$ are diagonal elements of the matrices

$$Q = X(X^T WX)^{-1}X^T \tag{14}$$

and

$$W = \frac{(d_i)^2}{k_{2i}} \tag{15}$$

The $k_{2i}$ and $k_{3i}$ are the variances of response values $y$ and the third-order cumulant[10] for the Bernoulli distribution, which are $\pi(1-\pi)$ and $\pi(1-\pi)(1-2\pi)$, respectively. The derivation of the cumulant generating function for the Bernoulli distribution is presented in Appendix A.1. The $d_i$ in $W$ is related to eqs 3 and 4 and the probit

$$d_i = \frac{d\pi}{d\theta} = f(x_i) \tag{16}$$

For a non-canonical link such as probit, the $k_{3i}$ and $k_{2i}$ are replaced with $d_i'$ and $d_i$ because relations $k_{2i} = \emptyset d_i$ and $k_{3i} = \emptyset^2 d_i'$ do not apply, where $\emptyset$ is a dispersion parameter. The $d_i'$ for the probit is

$$d_i' = \frac{d^2\pi}{d\theta^2} = -(x_i\beta)f(x_i\beta) \tag{17}$$

Further discussion of the non-canonical link is provided in Section 2.4. The $W$s for the logistic and probit binomials are, respectively, the following diagonal elements:

$$W = \text{diag}\{\pi_i(1-\pi_i)\} \tag{18}$$

and

$$W = \text{diag}\{\frac{f_i^2}{\pi_i(1-\pi_i)}\} \tag{19}$$

The denominator in eq 19 is the inverse of the variance of $y$ response values, multiplied by a square of the normal density function. The numerator comes from the non-canonical link shown in eq 4.

The BC method subtracts first-order bias from the parameter estimates, which were infinite for the CS and quasi-CS data. Thus, the BC method fails to eliminate the initial problem of infinite estimates.

### 2.3 BR Method Derivation

Firth suggested that the reduction of the first-order bias for the EFCL could be performed by modifying the score function, $U^*(\beta) = U(\beta) - I(\beta)b(\beta)$, where $I(\beta)$ is the information matrix, and $b(\beta)$ is the first-order bias in the maximum likelihood estimator, $\tilde{\beta}$.[3] In his paper, he explained that this method shifts the score function downward by applying the

information matrix, which is the gradient of the score function. Unlike the BC method, this method does not depend on the maximum likelihood estimate. The modified score function for the EFCL can be expressed as $U^* = U - X^T W\epsilon$, if we apply the first-order bias term in eq 12 to the information matrix, which will be explained in detail in Section 2.4 and Appendix A.2.

A general form of MMLE for the canonical link is[3]

$$U_t^* = U_t + \frac{1}{2\emptyset}\sum_i \left(\frac{k_{3i}}{k_{2i}}\right) h_i x_{ir} \qquad (t = 1, \ldots, p) \tag{20}$$

where $h_i$ is a diagonal element

$$H = W^{\frac{1}{2}} X (X^T W X)^{-1} X^T W^{\frac{1}{2}} \tag{21}$$

and $W$ is given in eq 15.

Firth also showed a relation between the MMLE and the PL functions for the EFCL for the first-order bias reduction expressed as $U(\beta_i)^* \equiv U(\beta_i) + \frac{1}{2}\text{trace}\left[I(\beta)^{-1}\left\{\frac{\sigma I(\beta)}{\sigma \beta_i}\right\}\right]$, where $I(\beta)^{-1}$ is an inverse of the information matrix evaluated at $\beta$. The solution of the above MMLE is the stationary points for $L^*(\beta)^* = L(\beta)|I(\beta)|^{\frac{1}{2}}$, where $|I(\beta)|^{\frac{1}{2}}$ is the Jeffreys invariant prior.[11] Therefore, the same BR result can be achieved for the EFCL by applying the PL method.[3]

The BR estimate is also consistent and asymptotically normal. One of the properties for the BR method for the EFCL is that it always has finite estimates, even though the traditional MLE has infinite values.[3]

## 2.4 BR Method Derivation for the Non-Canonical Model

A general link form of a modified score function for the non-canonical model may be expressed as[12]

$$U_t^* = U_t + \frac{1}{2}\sum_i h_i \frac{d_i'}{d_i} x_i \tag{22}$$

where $t$ indexes parameters 1,…,$p$, and $i$ is the number of sample sizes. The $d_i$, $d_i'$, and $h_i$ are given in eqs 16, 17, and 21. The difference between the canonical and the non-canonical models is that the latter involves the covariance between the first- and second-order derivatives of a likelihood function.[6]

## 3. ESTIMATION METHOD

### 3.1 Estimation Method with Newton Raphson Iteration

A first-order Taylor-series expansion for the modified score function about the true unknown parameters, $\beta_0$, may be expressed as[13]

$$0 = U^*(\tilde{\beta}) \approx U^*(\beta_0) - I(\beta_0)(\tilde{\beta} - \beta_0) \tag{23}$$

where $I(\beta_0)$ is an information matrix. From eq 23,

$$\tilde{\beta} \approx \beta_0 + \{I(\beta_0)\}^{-1} U^*(\beta_0) \tag{24}$$

and final parameter estimates can be obtained by updating initial estimates until the estimates converge to certain criteria. In this approach, the values for $\beta_0$ are replaced with initial starting values:

$$\beta_{s+1} = \beta_s + \{I^{-1}(\beta_s) U_s^*\} \tag{25}$$

If the relevant eqs 7, 16, 17, and 21 are used in eq 22 for the probit, the MMLE model can be written as

$$U_t^* = \sum_i \frac{f(x_i\beta)}{F(x_i\beta)(1-F(x_i\beta))}\{y_i - \frac{1}{2} h_i \frac{F(x_i\beta)(1-F(x_i\beta)(x_i\beta))}{f(x_i\beta)} - F(x_i\beta)\} x_{i,t} \quad t = 1,\ldots,p \tag{26}$$

Matrices for an information matrix and a modified score function can be formed from eqs 10 and 26, which would then be used for the iterative reweighted estimation process in eq 24. A variance-covariance matrix can be obtained from an inversion of the information matrix expressed in eq 10.

### 3.2 Initial Estimates for the Newton Raphson Procedure

Finding suitable initial estimates is a key factor in getting converged estimates in eq 24; otherwise, estimates may go out of range for a normal density. A cumulative normal distribution may not generate proper values in eq 26.

There are several ways to generate suitable initial estimates, such as by augmenting data for the two-response values (0 and 1) to make pseudo-overlapped data or by dividing large estimated values from the traditional MLE by a constant that does not generate extreme values in eq 26. Estimates from an ordinary regression can also be used if the GLM function is not available. Alternatively, the simplest way is to set all the initial parameters at 0, thus generating 0.5 and 0.40 for the cumulative normal distributions in eq 5 and the density

function in eq 8, respectively. The last method may be preferable over methods for other platforms because it does not require any calculations.

Comparisons of parameter estimations and variances between the BR and traditional MLE methods for the three different data sets that were shown in Figures 1–3 are provided in Tables 1–3. Comparisons of LD50 and CI values are provided in Table 4–6.

Table 1. Comparisons of Parameter Estimates and Standard Errors between BR and Traditional MLE Methods for Figure 1

| Parameter | Estimate | | Standard Error | |
|---|---|---|---|---|
| | BR | Traditional | BR | Traditional |
| Intercept | −6.95 | −102.07 | 3.74 | 95865 |
| Slope for Dose | 5.20 | 73.52 | 2.83 | 69511 |

Table 2. Comparisons of LD50 and CI Values between BR and Traditional MLE Methods from Table 1

| Method | LD50 | Standard Error | CI |
|---|---|---|---|
| BR | 21.68 | 0.10 | (13.68, 34.37) |
| Traditional MLE | 24.45 | 83.14 | $(-\infty^*, \infty)$ |

*∞ denotes infinity.

Table 3. Comparisons of Parameter Estimates and Standard Errors between BR and Traditional MLE Methods for Figure 2

| Parameter | Estimate | | Standard Error | |
|---|---|---|---|---|
| | BR | Traditional | BR | Traditional |
| Intercept | −5.64 | −60.03 | 3.62 | 12250 |
| Slope for Dose | 11.43 | 120.34 | 6.94 | 24381 |

Table 4. Comparisons of LD50 and CI Values between BR and Traditional MLE Methods from Table 3

| Method | LD50 | Standard Error | CI |
|---|---|---|---|
| BR | 3.12 | 0.05 | (2.54, 3.82) |
| Traditional MLE | 3.15 | 0.73 | (0.12, 83.22) |

Table 5. Comparisons of Parameter Estimates and Standard Errors between BR and Traditional MLE Methods for Figure 3

| Parameter | Estimate | | Standard Error | |
|---|---|---|---|---|
| | BR | Traditional | BR | Traditional |
| Intercept | −8.01 | −9.92 | 3.24 | 3.87 |
| Slope for Dose | 5.47 | 6.73 | 2.14 | 2.55 |

Table 6. Comparisons of LD50 and CI Values between BR and Traditional MLE Methods from Table 5

| Method | LD50 | Standard Error | CI |
|---|---|---|---|
| BR | 29.05 | 0.05 | (23.87, 35.34) |
| Traditional MLE | 29.74 | 0.04 | (25.26, 35.01) |

## 4. SUMMARY

The BR method is a variant of the BC method, and, unlike the latter, it does not require finite estimates in removing the first-order bias term in the asymptotic expansion of MLE. For the EFCL, the BR method results in the same parameter estimates as the PL method using the Jeffreys invariant prior. It also generates finite estimates, although the traditional MLE produces infinite estimates.

For CS or quasi-CS data, the BR method does not require prior information like Bayesian techniques, which may be difficult in many different situations. Like the traditional MLE method, the BR estimator has consistency and asymptotic normality. A variance-covariance matrix can be obtained from an inversion of the information matrix after the reweighted iterative procedure is finished.

Short descriptions of formula derivations are provided because a detailed discussion for complicated mathematical theory developments is not the focus of this paper.

R codes (Appendix B) are provided for executing the BR method, which uses 0 intercept and 0 slope for initial estimates, and the convergence criteria for the estimates are set to $1 \times 10^{-5}$ in the codes for calculating the differences between the two successive estimates. The written BR method is the same as that in the binomial-response GLM program in R, and the estimated values are matched between the two.

All the necessary formulas and procedures are provided to execute the BR method to estimate LD50 or ED50 for binomial distribution with probit link for CS and quasi-CS data. These formulas are useful in getting estimates for different computing platforms, which do not have a built-in module for MMLE of GLM. The only required computing capabilities are inversion and multiplication of a matrix with modules capable of generating values from a normal density function and a cumulative normal distribution function (eqs 5 and 8, respectively).

The BR method is recommended as an alternative to applying related probit slopes, which have been used historically at ECBC for CS and quasi-CS data. When probit slopes are unavailable and when sample sizes are small, the BR method may enable previously unattainable analyses.

## LITERATURE CITED


1. Albert, A.; Anderson J.A. On the Existence of Maximum Likelihood Estimates in Logistic Models. *Biometrika* **1984**, *71*(1), 1–10.

2. Silvapulle, M.J. On the Existence of Maximum Likelihood Estimates for the Binomial Response Models. *J. R. Statist. Soc. B* **1981,** *43*, 310–313.

3. Firth, D. Bias Reduction of Maximum Likelihood Estimates. *Boimetrika* **1993**, *80*(1), 27–38.

4. Clogg, C.C.; Rubbin, D.B.; Schenker, N.; Schultz, B.; Weidman, L. Multiple Imputation of Industry and Occupation Codes in Census Public-Use Samples Using Bayesian Logistic Regression. *JASA* **1991**, *86*(413), 68–78.

5. Gelman, A.; Jakulin, A.; Pittau, M.G.; Yu, S.S. A Weakly Informative Default Prior Distribution for Logistic and other Regression Model. *Ann. Appl. Stat.* **2008**, *2*, 1360–1383.

6. McCullagh, P.; Nelder, J.A. *Generalized Linear Model,* 2nd ed.; Chapman and Hall: London, 1989.

7. Johnston, J. *Econometric Methods,* 3rd ed.; McGraw-Hill Companies: New York, 1984.

8. McLachlan, G.J. A Note on Bias Correction in Maximum Likelihood Estimation with Logistic Discrimination. *Technometrics* **1980**, *22*, 621–627.

9. Cox, D.R.; Snell, E.J. *Analysis of Binary Data,* 2nd ed.; Chapman and Hall: London, 1970.

10. Kendall, M.G.; Stuart A. *The Advanced Theory of Statistics,* 3rd ed.; Griffin and Company, Ltd.: London, 1969; Vol. 1.

11. Jefferys, H. An Invariant Form for the Prior Probability in Estimation Problems. *Proc. R. Soc. Lond. A Math. Phys. Sci.* **1946**, *186*(1007), 453–461.

12. Kosmidis, I.; Firth, D. Bias Reduction in Exponential Family Nonlinear Models. *Biometrika* **2009**, *96* (4), 793–804.

13. Rohatgi, V.K. *Introduction to Probability Theory and Mathematical Statistics*; John Wiley and Sons, Inc.: New York, 1976.

Blank

## ACRONYMS AND ABBREVIATIONS

| | |
|---|---|
| BC | bias correction |
| BR | bias reduction |
| CI | confidence interval |
| CS | completely separated |
| DR | dose response |
| ECBC | U.S. Army Edgewood Chemical Biological Center |
| ED50 | $50^{th}$ percentile of the effective dose |
| EFCL | exponential family with canonical link |
| GLM | generalized linear model |
| HPAC | Hazard Prediction and Assessment Capability |
| LD50 | $50^{th}$ percentile of the lethal dose |
| MLE | maximum likelihood estimation |
| MMLE | modified maximum likelihood estimation |
| PL | penalized log-likelihood |
| quasi-CS | quasi-completely separated |
| VLSTrack | Vapor, Liquid, and Solid Tracking |

Blank

## APPENDIX A

## DERIVATION OF CUMULANT GENERATING AND MODIFIED MAXIMUM LIKELIHOOD ESTIMATION FUNCTIONS

### A.1 Derivation of the Cumulant Generating Function for Bernoulli Distribution

$$g'(t) = \frac{1}{\left(\left(\frac{1}{p}-1\right)e^{-t}+1\right)}. \quad \text{(A1)}$$

In eq A1, $k_1 = g'(0) = p$, $k_2 = g'(0) = p(1-p)$. It can be calculated from the recursive formula in eq A2

$$k_{n+1} = p(1-p)\frac{dk_n}{dp} \quad \text{(A2)}$$

### A.2 Derivation of Modified Maximum Likelihood Estimation (MMLE) Function

Because a general form of a score function for exponential form is[†] $U = X^t W D^{-1}(y-\pi)$, and the expected information matrix is expressed as $X^TWX$, the reduction of the first term is $I(\beta) \times \frac{b_1(\beta)}{n} = (X^TWX) \times (X^TWX)^{-1}X^TW \in$, where $W$ is derived from $W = \frac{(d_i)^2}{k_{2i}}$, and $\epsilon$ for a general form is[‡] $\epsilon_i = -\frac{1}{2}Q_{ii}\frac{d_i'}{d_i}$. The $Q$, $d_i$, and $d_i'$ are shown by the following functions: $Q = X(X^TWX)^{-1}X^T$; $d_i = \frac{d\pi}{d\theta} = f(x_i)$; and $d_i' = \frac{d^2\pi}{d\theta^2} = -(x_i\beta)f(x_i\beta)$, respectively. From eqs A1 and A2, the MMLE is expressed as $U_t^* = X^TWD^{-1}(y-\pi) - X^TW\epsilon$. This can be rewritten as $U_t^* = X^TWD^{-1}(y - D\epsilon - \pi)$. If $W$, $D$, and $\epsilon$ are used as substitutes in the score equation, its components can be restated as $U_t^* = \sum_i \frac{d_r}{k_{2,r}}(y_i + \frac{1}{2}h_i\frac{d_i'}{w_i} - \pi_i)x_{it}$.

[†] Kosmidis, I.; Firth, D. Bias Reduction in Exponential Family Nonlinear Models. *Biometrika* **2009**, *96* (4), 793–804.

[‡] McCullagh, P.; Nelder, J.A. *Generalized Linear Model*, 2nd ed.; Chapman and Hall: London, 1989.

Blank

## APPENDIX B

## R CODES USED FOR ESTIMATING AN INTERCEPT AND A SLOPE WITH THE BIAS-REDUCTION METHOD

```
function(data)
{
# This is finding an intercept and a slope for probit analysis for separated data for one covariate
'dose'.
# The solution is based on Ioannis Kosmidis's Ph.D. dissertation.*
# For initial estimates, 0 intercept and 0 slope are used.
# First, sort the dose in ascending order.
data<-data[order(data$dose),]
dose<-data$dose
log.dose<-log10(dose)
# Get a length of data.
data.dim<-dim(data)[1]
old.inter<-10000
old.sp<-10000
# initial estimates of an intercept and a slope
inter<-0
sp<-0

while (abs(old.inter-inter)>=0.000005 && abs(old.sp-sp)>=0.000005)
{
resp<-data$resp
# To add an intercept term
x.inter<-rep(1,data.dim)
x.sp<-log.dose
# This is from the first derivative
w<-(dnorm(inter+sp*x.sp)/(pnorm(inter+sp*x.sp)*(1-pnorm(inter+sp*x.sp))))
# This is from second derivative.
w1<-(dnorm(inter+sp*x.sp))^2/(pnorm(inter+sp*x.sp)*(1-pnorm(inter+sp*x.sp)))

library(Matrix)
# This is for H matrix
w1.diag<-Diagonal(data.dim,w1)
x.total<-cbind(x.inter,x.sp)
t.x.total<-t(x.total)
middle.invert<-solve(t.x.total%*%w1.diag%*%x.total)
h<-sqrt(w1.diag)%*%x.total%*%middle.invert%*%t.x.total%*%sqrt(w1.diag)
# Now pick up the diagonal elements of H matrix.
h.diag<-diag(h)
```

* Kosmidis, I. Bias Reduction in Exponential Family Nonlinear Models. Ph.D. Dissertation, University of Warwick, Coventry, U.K., 2007.

```
# This is for an intercept in the first derivative.
# ub1<-w*(resp-pnorm(inter+sp*x.sp)+h.diag*(0.5-pnorm(inter+sp*x.sp)))*x.inter
# This is from R score adjusted function ub1<-w*(resp-0.5*h.diag*(pnorm(inter+sp*x.sp)*(1-
pnorm(inter+sp*x.sp))*(inter+sp*x.sp)/dnorm(inter+sp*x.sp))-pnorm(inter+sp*x.sp))*x.inter
ub1.sum<-sum(ub1)

ub2<-w*(resp-0.5*h.diag*(pnorm(inter+sp*x.sp)*(1-
pnorm(inter+sp*x.sp))*(inter+sp*x.sp)/dnorm(inter+sp*x.sp))-pnorm(inter+sp*x.sp))*x.sp
ub2.sum<-sum(ub2)
ub.mat<-matrix(c(ub1.sum,ub2.sum),nrow=2)

# Generate a hessian matrix.
he11<-sum(w1*x.inter)
he12<-sum(w1*x.inter*x.sp)
he21<-sum(w1*x.inter*x.sp)
he22<-sum(w1*x.sp^2)
he.mat<-matrix(c(he11,he12,he21,he22),nrow=2,byrow=T)
new.mat.solve<-solve(he.mat)%*%ub.mat
old.mat<-matrix(c(inter,sp),nrow=2)

old.inter<-old.mat[1,1]
old.sp<-old.mat[2,1]
# Upgrade to new estimates.
new.mat<-old.mat+new.mat.solve
inter<-new.mat[1,1]
sp<-new.mat[2,1]
}
# Final estimates
est<<-matrix(c(inter,sp),byrow=T)
#new.mat<<-new.mat
va.final<-solve(he.mat)

# To retrieve the final variances matrix,
va.final<<-va.final
}
```

## DISTRIBUTION LIST

The following individuals and organizations were provided with one Adobe portable document format (pdf) electronic version of this report:

U.S. Army Edgewood Chemical
Biological Center (ECBC)
RDCB-DRI-M
ATTN: Park, K.
Lagan, S.J.
Kierzewski, M.

Defense Threat Reduction Agency
DTRA/RD-CBD T
ATTN: Ward, T.
J9-CBS
ATTN: Moore, E.

Department of Homeland Security
DHS ORD CSAC
ATTN: Famini, G.

G-3 History Office
U.S. Army RDECOM
ATTN: Smart, J.

ECBC Technical Library
RDCB-DRB-BL
ATTN: Foppiano, S.
Stein, J.

Office of the Chief Counsel
AMSRD-CC
ATTN: Upchurch, V.

Defense Technical Information Center
ATTN: DTIC OA

ECBC Rock Island
RDCB-DES
ATTN: Lee, K.


EDGEWOOD CHEMICAL BIOLOGICAL CENTER
A U.S. Army RDECOM Laboratory